PATENT APPLICATION
VIXS.0100020

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

FILING OF A UNITED STATES PATENT APPLICATION

ADAPTIVE BANDWIDTH FOOTPRINT MATCHING FOR MULTIPLE COMPRESSED VIDEO STREAMS IN A FIXED BANDWIDTH NETWORK

INVENTORS:

INDRA LAKSONO
138 Old Hill Street
Richmond Hill, Ontario
Canada L4C 9Z7

ATTORNEY OF RECORD
J. GUSTAV LARSON

SIMON, GALASSO & FRANTZ, PLC
P.O. Box 26503
Austin, TX 78755-0503
PHONE (512) 372-8240
FAX (512) 372-8247

Express Mail Label No. EL855710995US

Date of Deposit: 3-30-01

I hereby certify that this paper is being deposited with the U.S. Postal Service "Express Mail Post Office to Addresses" service under 37 C.F.R. Section 1.10 on the 'Date of Deposit', indicated above, and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Name of Depositor: Terri Alloway
(print or type)

Signature: Terri Alloway

ADAPTIVE BANDWIDTH FOOTPRINT MATCHING FOR MULTIPLE COMPRESSED VIDEO STREAMS IN A FIXED BANDWIDTH NETWORK

FIELD OF THE DISCLOSURE

The present invention relates generally to media data transmission and more particularly to reducing bandwidth overload.

BACKGROUND

A number of media playback systems use continuous media streams, such as video image streams, to output media content. However, some continuous media streams in their raw form often require high transmission rates, or bandwidth, for effective and/or timely transmission. In many cases, the cost and/or effort of providing the required transmission rate is prohibitive. This transmission rate problem is often solved by compression schemes that take advantage of the continuity in content to create highly packed data. Compression methods such Motion Picture Experts Group (MPEG) methods and its variants for video are well known in the art. MPEG and similar variants use motion estimation of blocks of images between frames to perform this compression. With extremely high resolutions, such as the resolution of 1920x1080i used in high definition television (HDTV), the data transmission rate of such a video image stream will be very high even after compression.

One problem posed by such a high data transmission rate is data storage. Recording or saving high resolution video image streams for any reasonable length of time requires considerably large amounts of storage that can be prohibitively expensive. Another problem presented by a high data transmission rate is that many output devices are incapable of handling the transmission. For example, display systems that can be used to view video image streams having a lower resolution may not be capable of displaying such a high resolution. Yet another problem is the limitations on continuous media streaming in systems with a fixed bandwidth or capacity. For example, in a local area network with multiple receiving/output devices, such a network will often have a fixed bandwidth or capacity, and hence be physically and/or logistically incapable of simultaneously supporting multiple receiving/output devices.

Given the limitations, as discussed, it is apparent that a method and/or system that overcome at least some of these limitations would be advantageous.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a state machine diagram illustrating an Adaptive Bandwidth Footprint Matching implementation according to at least one embodiment of the present invention;

FIG. 2 is a system diagram illustrating a server system for implementing Adaptive Bandwidth Footprint Matching according to at least one embodiment of the present invention;

FIG. 3 is a block diagram illustrating components of a gateway media server according to at least one embodiment of the present invention; and

FIG. 4 is a block diagram illustrating components of a receiver client unit according to at least one embodiment of the present invention.

DETAILED DESCRIPTION OF THE FIGURES

In accordance with at least one embodiment of the present invention, a display data is received. It is determined if a predetermined criteria is met by a first representation of the display data, wherein the first representation of the display data includes a first plurality of display streams to be transmitted to a second plurality of display devices. A first display stream of the first plurality of display streams is compressed in a first manner when it is determined that the first representation of the display does not meet the predetermined criteria. An advantage of the present invention is that networks for broadcasting of media streams are implemented more efficiently. Another advantage of the present invention is that multiple media streams may be transmitted to multiple users on a fixed bandwidth network by managing degradation in transmission quality.

FIGS. 1-4 illustrate a system and a method for transmission of multiple data streams in a network capable of supporting a fixed bandwidth. The system includes a central gateway media server and a plurality of client receiver units. The input data streams arrive from an external source, such as a satellite television transmission, or physical head end, and are transmitted to the client receiver units in a compressed format. The data streams can include display data, graphics data,

digital data, analog data, multimedia data, and the like. An Adaptive Bandwidth Footprint Matching state machine on the gateway media server detects if the network bandwidth is close to saturation. The start time of each unit of media for each stream is matched against the estimated transmission time for that unit. When any one actual transmission time exceeds its estimated transmission time by a predetermined threshold, the network is deemed to be close to saturation, or already saturated, and the state machine will execute a process of selecting at least one stream as a target for lowering total bandwidth usage. Once the target stream associated with a client receiver unit is chosen, the target stream is modified to transmit less data, which may result in a lower data transmission rate. For example, a decrease in the data to be transmitted can be accomplished by a gradual escalation of the degree of data compression performed on the target stream, thereby reducing the resolution of the target stream. If escalation of the degree of data compression alone does not adequately reduce the data to be transmitted to prevent bandwidth saturation, the resolution of the target stream can also be reduced. For example, if the target stream is a video stream, the frame size could be scaled down, reducing the amount of data per frame, and thereby reducing the data transmission rate.

Referring now to FIG. 1, a state machine diagram of an Adaptive Bandwidth Footprint Matching (ABFM) method with three kinds of degradation is illustrated according to at least one embodiment of the present invention, where degradation is used to reduce amount of data and/or the data rate associated with a given data stream. Although the following discussion makes use of video streams for ease of illustration, other data formats, such as audio, display, analog, digital, multimedia, and the like, may be used in accordance with various embodiments. In steady state 100, each video stream of a plurality of video streams is operating within acceptable parameters. In at least one embodiment, a video stream is determined to be acceptably operating when the transmission of a frame of video data is transmitted without exceeding a maximum allowed delay time. For example, digital video streams such as MPEG often have time stamp information embedded within the stream. In addition to the start time T0 (when the frame was successfully transmitted) of the first frame in a sequence of frames with a fixed interframe time, this time stamp information, including the known interframe time (which is fixed for a sequence of frames), can be used to calculate the estimated times of each frame as they arrive. For example, in one embodiment, the estimated time of frame transmission completion T'j for frame N is calculated as $T'j(N) = T0 + N*D$, where D is the interframe time. In this case, if the estimated times for transmission of the start

of each frame of a stream j is within acceptable limits and has not exceed a maximum allowed delay Dj, stream j can be considered as operating within acceptable parameters. The acceptable parameters may be set by an administrator, determined empirically, and the like.

The desired tolerance Dj (or maximum acceptable delay time) can be calculated using a variety of methods. In one embodiment, the method used is to take into consideration the buffering size of each client receiver unit, and ensure that the client receiver unit will not run out of media content to decode. A typical formula to calculate Dj is to take the size of the buffer and estimate a lower bound (in units of time) to consume, or fill, the buffer. As it is often desirable to keep the buffer of each client receiver unit as full as possible, a typical Dj will be calculated as Dj = Tj(estimate)/2. Where Tj(estimate) is the estimated lower time bound to completely consume the input buffer of a receiver unit associated with stream j. Alternately, instead of using 1/2 of Tj(estimate), a more aggressive approach would be to use 3/4 of Tj(estimate), and a more conservative approach might take 1/3 of Tj(estimate). In cases where Tj(estimate) is small for receiver devices that are incapable of providing considerable buffer space, a conservative approach may be more appropriate. In one embodiment, Tj(estimate) is obtained by taking observed peak (highest) data rate (in bytes/second) of stream j and the smallest size of the buffers (in bytes) of all the devices receiving stream j. In this case, Tj(estimate) can be evaluated as Bp/Rp, where Bp is the receive buffer size of device p and Rp is the peak data rate of stream j associated with device p, where device p receives stream j and has the smallest receive buffer. Alternately, Rp can be associated with any value between the mean (average) and the peak. In one embodiment, the peak data rate (Rp) can be based on the largest compressed frame. If the receiving client unit does not have enough buffering capability for at least one compressed frame then it is unlikely to be able to display the video smoothly without dropping frames.

At the commencement of each unit of media, such as a frame of video, the ABFM state machine transitions to state 110. In state 110, the actual transmit time Tj (the actual time of frame transmission completion) is compared against the estimated transmit time T'j (the expected time of frame transmission completion) at the start of each frame of stream j. In one embodiment, if the actual time of frame transmission completion exceeds the estimated time by less than the desired tolerance Dj (i.e. Tj - T'j < Dj), the ABFM state machine returns to steady state 100. Otherwise, if the actual transmit time exceeds the estimated time by at least desired tolerance Dj (i.e. Tj -T'j >=

Dj)), the ABFM state machine enters state 120.

In state 120, a victim stream v is selected from the plurality of video streams. In one embodiment, victim stream v is selected using a predetermined selection method, such as by round robin selection where each video stream is chosen in turn. In another embodiment, the victim stream v is selected based on a fixed priority scheme where lower priority streams are always selected before any higher priority scheme. In yet another embodiment, the victim stream v is selected based on a weighted priority scheme where the stream having the greatest amount of data and/or the priority of each stream plays a role in its probability of selection.

Regardless of the method of selecting a victim stream v, in one embodiment, each stream j has a count, herein referred to as A(j), that refers to the current degradation value of the modified stream of stream j. In this case, the current degradation value of victim stream v, A(v), is evaluated in state 120. If A(v) is 0, in one embodiment, the one or more quantization factors of the reencoding process for the victim stream v are changed in state 130, thus resulting in a decreased in the amount of data transmitted in victim stream v. In one embodiment, the quantization factors are increased resulting in a decrease in the amount of data transmitted in victim stream v. For example, the MPEG algorithm uses quantization factors to reduce the amount of data by reducing the precision of the transmitted video stream. MPEG relies on quantization of matrices of picture elements (pixels) or differences in values of pixels to obtain as many zero elements as possible. The higher the quantization factors, the more zero elements produced. Using algorithms such as run-length encoding, video streams (or their associated matrices) containing more zeros can be more highly compressed than video streams having fewer zeros.

For example, the MPEG algorithm for compression of a video stream has a stage in the algorithm for a discrete cosine transform (DCT), a special type of a Fourier Transform. The DCT is used to transform blocks of pixels in the time domain to the frequency domain. As a result of this transformation, the elements in the frequency domain, post-DCT, that are closest to the top left element of the resulting matrix with indices (0,0) are weighted more heavily compared to elements at the bottom right of the matrix. If the matrix in the frequency domain were to use less precision to represent the elements in the lower right half of the matrix of elements, the smaller values in the lower right half will get converted to zero if they are below a threshold based on a quantization

factor. Dividing each element by a quantization factor is one method utilized to produce more zero elements. MPEG and related algorithms often apply larger quantization values to decrease the precision of the matrices in the frequency domain, resulting in more zero elements, and hence a decrease the data transmission rate.

After the reducing the data transmission of the victim stream v by modifying the quantization factor (state 130), in one embodiment, the ABFM state machine transitions to state 160, where the degradation value A(v) is increased by one and then a modulus of 3 is applied, i.e. A(v)current = (A(v)previous + 1) mod 3. As a result, the value of A(v) can cycle from 0 to 2. Since A(v) was previously determined to be 1 in state 120, the new A(v) value would be 1 (0+1 mod 3). After modifying the degradation value A(v) for victim stream v in state 160, the ABFM state machine transitions back to state 100.

If A(v) is determined to be 1 for victim stream v in state 120, the ABFM state machine enters state 140. In one embodiment, the height of the reencoded data stream is reduced by a predetermined amount, ½ for example, in state 140, resulting in a decreased amount of data to be transmitted. One method used to scale blocks of pixels by half is to blend and average pixels. Another method used is to drop every other pixel. In cases where the video stream is interlaced, halving the height can be achieved by dropping alternate fields, such as dropping all of the odd horizontal display rows or all of the even horizontal display rows. It will be appreciated that in some formats, particularly those in the National Television System Committee (NTSC) and the Advanced Television System Committee (ATSC) formats, video streams are interlaced where the even horizontal display rows for an entire frame are displayed first and then the odd horizontal display rows are displayed next. In other embodiments, the height of the reencoded data stream is reduced by a factor other than a half, such as 1/3, using similar methods as appropriate.

After the reducing the data transmission of the victim stream v by reducing the resolution of the stream (state 140), in one embodiment, the degradation value A(v) is modified in state 160, as discussed previously. The resulting value for A(v) is 2 (1+1 mod 3). After modifying the degradation value A(v) for victim stream v in state 160, the ABFM state machine transitions back to state 100.

If A(v) is determined to be 2 for victim stream v in state 120, the ABFM state machine enters

state 150. In one embodiment, the width of the reencoded data stream is reduced by a predetermined amount in state 150 using methods similar to those discussed previously with reference to state 140, such as dropping every other pixel. It will be appreciated that for a same reduction factor, the reduction methods of state 140 or state 150 are interchangeable. In cases where the victim stream v is interlaced, halving the height before the width is generally more appropriate as it is more efficient to completely skip alternating fields, saving substantial processing requirements.

After the reducing the data transmission of the victim stream v by reducing the resolution of the stream (state 150), in one embodiment, the degradation value A(v) is modified in state 160, as discussed previously. The resulting value for A(v) is 0 (2+1 mod 3). After modifying the degradation value A(v) for victim stream v in state 160, the ABFM state machine transitions back to state 100.

In one embodiment, as a result of the cycling between 0 to 2 of the degradation value A(v) for a victim stream v, the ABFM state machine cycles through three different kinds of degradation of the resolution and/or the precision of victim stream v each time it is selected for degradation in state 120. Although an ABFM state machine utilizing three kinds of data degradation has been discussed, in other embodiments, fewer or more steps of data degradation may be used according to the present invention. For example, in one embodiment, an ABFM state machine utilizes multiple step degradation involving more than one state of changing of quantization factors. It will also be appreciated that scaling factors of width and height other than 1/2 (e.g.: 3/4) may be used. For example, in one embodiment, the amount by which the resolution and/or precision of a victim stream v is based on the degree to which the actual frame transmission completion time for a frame of video in victim stream v exceeds the estimated frame transmission completion time. For example, if the actual frame transmission completion time is 10% greater than the estimated frame transmission completion time, then the resolution of victim stream v could be scaled down by 10%, thereby causing the actual frame transmission completion time to likely come closer to the estimated frame transmission completion time.

Referring next to FIG. 2, Adaptive Bandwidth Footprint Matching (ABFM) server system 205 is illustrated according to at least one embodiment of the present invention. Data streams, such as video data, display data, graphics data, MPEG data, and the like, are input to gateway media

server 210. In one embodiment, two main inputs sources are used by gateway media server 210. One input is wide area network (WAN) connection 200 to provide high speed Internet access. The other input is a source of media streams, such as satellite television (using satellite dish 201) or cable television. In other embodiments, other input sources can be used, such as a local area network (LAN). WAN connection 200 and/or other used input sources which can include a network comprised of cable, twisted pair wires, fiber optic cable, a wireless radio frequency net work, and the like.

Gateway media server 210, in one embodiment, accepts one or more input data streams, such as digital video or display data, from satellite dish 201 and/or WAN 200. Each input data stream can include a plurality of multiplexed channels, such as MPEG data channels. Gateway media server 210 broadcasts the data streams and/or channels over a common medium (local data network 220) to one or more receiving client units, such as laptop 230, computer 240, or viewing unit 250. In one embodiment, there is a one-to-one correspondence between the number of data channels input to gateway media server 210 and the number of client receiver units to receive output data channels or streams. In another embodiment, there are fewer data channels or streams than there are receiver client units. In this case, two or more client receiver units may need to share one or more data channels or streams. Local data network 220 can include a local area network, a wide area network, a bus, a serial connection, and the like. Local data network 220 may be constructed using cable, twisted pair wire, fiber optic cable, etc. During broadcast of the data streams to the receiving client units, gateway media server 210, in one embodiment, applies the ABFM algorithm, as discussed previously with reference to FIG. 1, to manage the network traffic to assure consistent and sustained delivery within acceptable parameters, thereby allowing users to view the data stream seamlessly.

In at least one embodiment, the ABFM algorithm is utilized by gateway media server 210 to attempt to ensure that a representation of the display data meets a predetermined criteria. For example, gateway media server 210 may transmit the display data to receiver client units, where a video sequence displayed on the receiver client units is a representation of the displayed data. If the video sequence is simultaneously displayed properly in real time (the predetermined criteria) on a number receiver client units, gateway media server 210 may not need to take further action. Else if the video sequence is choppy, is not synchronized, is delayed, or is not received by all the designated receiver client units, the representation of the display data does not meet the predetermined criteria,

and gateway media server 210, in one embodiment, uses the ABFM method previously discussed to modify one or more of the data streams of display data to improve the display of the video sequence.

Sub A'

As discussed previously, in at least one embodiment, an ABFM algorithm is implemented to maintain the data transmission rate of ABFM server system 205 within a fixed bandwidth. In one embodiment, the bandwidth of ABFM server system 205 is fixed by the maximum bandwidth of the transmission medium (local data network 225) between gateway media server 210 and the client receiver units (laptop 230, computer 240, or viewing unit 250). For example, if local data network is a local area network having a maximum transmission rate of 1 megabit per second, the bandwidth of ABFM server system 205 may be fixed at a maximum of 1 megabit per second. Alternately, in another embodiment, the bandwidth of ABFM server system 205 could be a predetermined portion of the available bandwidth of the transmission medium (local data network 225). For example, if there are four ABFM server systems 205 connected to local data network 225 having a maximum transmission rate of 1 megabit per second, each ABFM server systems 205 could be predetermined to have a fixed bandwidth of 0.25 megabits per second (one fourth of the maximum available transmission rate).

Although the transmission medium between gateway media server 210 and client receiver units is often the factor which limits or fixes the bandwidth of ABFM server system 205, in one embodiment, the bandwidth of ABFM server system 205 is fixed by the rate at which gateway media server 205 is able to input one or more data streams, compress one or more of the data streams, and output the compressed (and uncompressed) data streams or channels to the client receiver units. For example, if gateway media server 205 can only process 1 megabits of data per second, but local data network 225 has a transmission rate of 10 megabits per second, the bandwidth of ABFM server system 205 may be limited to only 1 megabits per second, even though local data network 225 can transmit at a higher transmission rate. It will be appreciated that the bandwidth of ABFM server system 205 could be limited by other factors without departing from the spirit or the scope of the present invention.

Referring to FIG. 3, gateway media server 210 is illustrated in greater detail according to at least one embodiment of the present invention. Input media streams enter the system via digital tuner demultiplexors (DEMUX) 330, from which the appropriate streams are sent to ABFM

transcoder controller circuit 350. ABFM transcoder controller circuit 350, in one embodiment, includes one or more stream parsing processors 360 that perform the higher level tasks of digital media decoding, such as video decoding. Stream parsing processors 360 drive a series of media transcoding vector processors 390 that perform the low level media transcoding tasks. The intermediate and final results of the decoding and transcoding are stored in the device memory, such as dynamic random access memory (DRAM) 380. The final compressed transcoded data, in one embodiment, is transmitted according to a direct memory access (DMA) method via external system input/output (IO) bus 320 past north bridge 305 into the host memory (host DRAM 310). Processor 300, using a timer driven dispatcher, at an appropriate time, will route the final compressed transcoded data stored in host DRAM 310 to network interface controller 395, which then routes the data to local area network (LAN) 399.

Referring next to FIG. 4, receiver client unit 401 is illustrated according to at least one embodiment of the present invention. Receiver client unit 401 can include devices capable of receiving and/or displaying media streams, such laptop 230, computer 240, and viewing unit 250 (FIG. 2). The final compressed transcoded data stream as discussed with reference to FIG. 3 is transmitted to network interface controller 400 via LAN 399. The data stream is then sent to media decoder/renderer 420 via IO connect 410. IO connect 410 can include any IO connection method, such as a bus or a serial connection. Media decoder/renderer 420, in one embodiment, includes embedded DRAM 430 which can be used as an intermediate storage area to store the decoded data. In cases where the decoded data does not fit within embedded DRAM 430, media decoder/renderer 420 further includes DRAM 440 which is larger than embedded DRAM 430. As the compressed data is decoded, it is transmitted to receiver client IO bus 490 and eventually is picked up by the receiver client unit's host processor (not shown). In one embodiment, the host processor controls video decoder/renderer 420 directly and actively reads the rendered data. In other embodiments, the functions of video decoder/renderer 420 are performed on the host via a software application. In cases where the host processor is incapable of such decoding tasks, video decoder/renderer 420 performs part of or all of the decoding tasks.

One implementation of the invention is as sets of computer readable instructions resident in the random access memory of one or more processing systems configured generally as described in FIGS. 1-4. Until required by the processing system, the set of instructions may be stored in another

computer readable memory, for example, in a hard disk drive or in a removable memory such as an optical disk for eventual use in a CD drive or DVD drive or a floppy disk for eventual use in a floppy disk drive. Further, the set of instructions can be stored in the memory of another image processing system and transmitted over a local area network or a wide area network, such as the Internet, where the transmitted signal could be a signal propagated through a medium such as an ISDN line, or the signal may be propagated through an air medium and received by a local satellite to be transferred to the processing system. Such a signal may be a composite signal comprising a carrier signal, and contained within the carrier signal is the desired information containing at least one computer program instruction implementing the invention, and may be downloaded as such when desired by the user. One skilled in the art would appreciate that the physical storage and/or transfer of the sets of instructions physically changes the medium upon which it is stored electrically, magnetically, or chemically so that the medium carries computer readable information.

In the preceding detailed description of the figures, reference has been made to the accompanying drawings which form a part thereof, and in which is shown by way of illustration specific preferred embodiments in which the invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention, and it is to be understood that other embodiments may be utilized and that logical, mechanical, chemical and electrical changes may be made without departing from the spirit or scope of the invention. To avoid detail not necessary to enable those skilled in the art to practice the invention, the description may omit certain information known to those skilled in the art. Furthermore, many other varied embodiments that incorporate the teachings of the invention may be easily constructed by those skilled in the art. Accordingly, the present invention is not intended to be limited to the specific form set forth herein, but on the contrary, it is intended to cover such alternatives, modifications, and equivalents, as can be reasonably included within the spirit and scope of the invention. The preceding detailed description is, therefore, not to be taken in a limiting sense, and the scope of the present invention is defined only by the appended claims.